

# UC Davis

## UC Davis Previously Published Works

### Title

Beyond the MHC: A canine model of dermatomyositis shows a complex pattern of genetic risk involving novel loci.

### Permalink

<https://escholarship.org/uc/item/3rd6s169>

### Journal

PLoS genetics, 13(2)

### ISSN

1553-7390

### Authors

Evans, Jacquelyn M  
Noorai, Rooksana E  
Tsai, Kate L  
et al.

### Publication Date

2017-02-01

### DOI

10.1371/journal.pgen.1006604

Peer reviewed

RESEARCH ARTICLE

# Beyond the MHC: A canine model of dermatomyositis shows a complex pattern of genetic risk involving novel loci

Jacquelyn M. Evans<sup>1</sup>, Rooksana E. Noorai<sup>1,2</sup>, Kate L. Tsai<sup>1</sup>, Alison N. Starr-Moss<sup>1</sup>, Cody M. Hill<sup>1</sup>, Kendall J. Anderson<sup>1</sup>, Thomas R. Famula<sup>3</sup>, Leigh Anne Clark<sup>1\*</sup>

**1** Department of Genetics and Biochemistry, Clemson University, Clemson, South Carolina, United States of America, **2** Genomics and Computational Laboratory, Clemson University, Clemson, South Carolina, United States of America, **3** Department of Animal Science, University of California, Davis, California, United States of America

\* [lclark4@clemson.edu](mailto:lclark4@clemson.edu)



## OPEN ACCESS

**Citation:** Evans JM, Noorai RE, Tsai KL, Starr-Moss AN, Hill CM, Anderson KJ, et al. (2017) Beyond the MHC: A canine model of dermatomyositis shows a complex pattern of genetic risk involving novel loci. *PLoS Genet* 13(2): e1006604. doi:10.1371/journal.pgen.1006604

**Editor:** Gregory S. Barsh, Stanford University School of Medicine, UNITED STATES

**Received:** December 15, 2016

**Accepted:** January 24, 2017

**Published:** February 3, 2017

**Copyright:** © 2017 Evans et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data availability statement:** SNP data are available in dbSNP under BioProject number PRJNA338128. Whole genome and transcriptome data were deposited in SRA genomes under accession number SRP081080.

**Funding:** Research reported in this publication was supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health under Award Number R15AR062868 ([www.niams.nih.gov](http://www.niams.nih.gov)), Collie Health Foundation ([www.colliehealth.org](http://www.colliehealth.org)), American

## Abstract

Juvenile dermatomyositis (JDM) is a chronic inflammatory myopathy and vasculopathy driven by genetic and environmental influences. Here, we investigated the genetic underpinnings of an analogous, spontaneous disease of dogs also termed dermatomyositis (DMS). As in JDM, we observed a significant association with a haplotype of the major histocompatibility complex (MHC) (*DLA-DRB1\*002:01/-DQA1\*009:01/-DQB1\*001:01*), particularly in homozygosity ( $P$ -val = 0.0001). However, the high incidence of the haplotype among healthy dogs indicated that additional genetic risk factors are likely involved in disease progression. We conducted genome-wide association studies in two modern breeds having common ancestry and detected strong associations with novel loci on canine chromosomes 10 ( $P$ -val =  $2.3 \times 10^{-12}$ ) and 31 ( $P$ -val =  $3.95 \times 10^{-8}$ ). Through whole genome resequencing, we identified primary candidate polymorphisms in conserved regions of *PAN2* (encoding p.Arg492Cys) and *MAP3K7CL* (c.383\_392ACTCCACAAA>GACT) on chromosomes 10 and 31, respectively. Analyses of these polymorphisms and the MHC haplotypes revealed that nine of 27 genotypic combinations confer high or moderate probability of disease and explain 93% of cases studied. The pattern of disease risk across *PAN2* and *MAP3K7CL* genotypes provided clear evidence for a significant epistatic foundation for this disease, a risk further impacted by MHC haplotypes. We also observed a genotype-phenotype correlation wherein an earlier age of onset is correlated with an increased number of risk alleles at *PAN2* and *MAP3K7CL*. High frequencies of multiple genetic risk factors are unique to affected breeds and likely arose coincident with artificial selection for desirable phenotypes. Described herein is the first three-locus association with a complex canine disease and two novel loci that provide targets for exploration in JDM and related immunological dysfunction.

Shetland Sheepdog Association ([www.assa.org](http://www.assa.org)), and AKC Canine Health Foundation ([www.akcchf.org](http://www.akcchf.org)). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

## Author summary

Juvenile dermatomyositis (JDM) is an autoimmune disease of the skin and muscle influenced by both genetic and environmental components. Although genes independent of the MHC are thought to contribute to disease pathogenesis, their identification has been complicated by a paucity of biological samples, disease heterogeneity, and genetically diverse subjects. A naturally occurring inflammatory disease of domestic dogs, also termed dermatomyositis (DMS), is analogous to JDM and is the only animal model available for genetic study. We first determined that, as in JDM, a particular MHC haplotype confers susceptibility to DMS. Capitalizing on the genetic isolation of dog breeds and extremely low MHC diversity within affected breeds, we used multibreed genome-wide association studies to identify two novel loci. Through genome resequencing and additional genotyping, we uncovered highly associated polymorphisms in conserved positions of *PAN2* and *MAP3K7CL*. Analysis of three-locus genotypes revealed uniquely high frequencies among affected breeds and nine allelic combinations that confer moderate or high risk for DMS. The pattern of disease probability illustrates the presence of gene-gene interaction, as well as an inverse correlation between age of onset and number of risk alleles. This study highlights the utility of canine models for mapping susceptibility loci in complex diseases and detecting patterns of genetic interactions obscured in diverse human populations.

## Introduction

Juvenile dermatomyositis (JDM) is an autoimmune vasculopathy that causes a characteristic skin rash (heliotrope rash across the eyelids and Gottron's papules on the bony prominences) and proximal muscle weakness [1]. It is the most frequently diagnosed childhood inflammatory myopathy, comprising 80% of all cases [1] and affecting 3.2 in every million children between the ages of 2 and 17 within the USA [2]. Prognosis is positively correlated with early diagnosis and swift treatment with corticosteroids and/or immunosuppressants [1,3]. While treatment of JDM is much improved overall, many children still suffer from chronic flare-ups [1].

Though the etiology is unknown, JDM is thought to be triggered by exposure to a virus or other environmental agent. Manlhiot et al. [4] reported that 71% of JDM patients had a clinical history consistent with infection prior to symptoms. Investigations into the class II major histocompatibility complex (MHC), *TNF*, and *IL1* identified several susceptibility and protective alleles, but their collective contribution to pathogenesis is poorly understood [5–8]. Recent genome-wide association studies (GWASs) to identify additional susceptibility loci in JDM confirmed a strong association with the MHC but failed to detect novel major risk factors, likely because of a paucity of biological samples and genetically heterogeneous populations [9,10].

In domestic dogs, an inflammatory vasculopathy of the skin and muscle, also termed dermatomyositis (DMS), is clinically, histologically, and immunologically similar to JDM and provides the only animal model available to study genetic risk factors [11–16]. The earliest clinical signs of DMS are crusting and scaling on the face, ears, tail tip, and/or across the bony prominences of the limbs and feet [17–19] (S1 Fig). Alopecia and more extensive skin lesions may develop over time, resulting in dermal scarring associated with erythema and mottled pigmentation [17,19]. Lesions persist for weeks to months, and may or may not chronically recur

[17]. Muscle wasting manifests as atrophy of the head musculature; difficulty eating, drinking, and swallowing; and an atypical, high-stepping gait [17,19].

Similar to JDM, DMS is an immune-mediated disease [13,18,20] that typically develops following an environmental trigger, such as vaccination or viral infection, and is exacerbated by subsequent stressors like exposure to UV light [13,17,21,22]. Anecdotal reports indicate that rabies vaccination, parvovirus infection, owner surrender, or mistreatment commonly precede disease onset. Consistent with an environmental trigger, age at onset is variable with many cases occurring between seven weeks and six months of age, but others not developing until well into adulthood [17–19,23].

DMS is diagnosed almost exclusively in the genetically [24] and phenotypically similar collie and Shetland sheepdog breeds, suggesting the presence of a strong heritable component(s) arising from ancestors common to both breeds. A 1980s study of disease transmission in the collie eliminated simple Mendelian modes of inheritance [14]. In two test matings, an affected male collie sired litters from an affected collie and a healthy Labrador retriever. All six collie puppies were affected with variable degrees of severity, while three of the four mixed breed puppies developed milder forms of DMS. Retrospective pedigree analyses of the collie sire and dam showed a complete absence of affected ancestors [14].

The availability of a naturally-occurring canine model provides a new opportunity for the identification of genetic risk factors of JDM. The conserved genomic backgrounds of genetically isolated dog breeds have enabled detection of risk loci in complex diseases that are often obscured by heterogeneity within human cohorts [25–30]. Here, we evaluated class II MHC haplotypes, performed multibreed GWASs, and generated whole genome and transcriptome sequencing data to dissect the genetic basis of DMS. We uncovered common polymorphisms of the MHC and two novel loci that are strongly associated with DMS, as well as patterns of allelic inheritance that explain 93% of cases studied. A genetic test is now available to determine the likelihood of a dog developing DMS and to facilitate breeding decisions that avoid progeny having high-risk genotypes.

## Results and discussion

### Association of a major histocompatibility complex haplotype

Given the involvement of MHC genes in JDM, we first determined alleles of the highly polymorphic canine MHC class II dog leukocyte antigen (DLA) genes: *DLA-DRB1*, *-DQA1*, and *-DQB1*. Two locus (*DLA-DRB1* and *-DQB1*) and three locus (*DLA-DRB1*, *-DQA1*, and *-DQB1*) haplotypes were first generated for 50 collies and 117 Shetland sheepdogs, respectively. Because all observed haplotypes contained a unique *DLA-DRB1* allele, the 355 remaining dogs were genotyped for this locus only and the haplotype was inferred (Table 1).

We observed remarkably low DLA diversity among collies, with only three haplotypes present in 225 collies worldwide. This lack of heterogeneity precluded detection of associations with DMS, as 91% of collies were homozygous for the haplotype *DLA-DRB1\*002:01/-DQA1\*009:01/-DQB1\*001:01*. In 297 Shetland sheepdogs, we identified two predominant haplotypes, *002:01/009:01/001:01* and *023:01/003:01/005:01*. The former was over-represented among cases ( $P$ -val = 0.0010, OR = 2.20), primarily because of increased homozygosity ( $P$ -val = 0.0001, OR = 2.98). We therefore conclude that *002:01/009:01/001:01* is a risk factor for DMS and that homozygosity confers increased susceptibility. Under the assumption that the causal alleles derive from an ancestor common to both breeds, we extrapolate the observed DLA risk to collies. The high frequency of the DLA risk haplotype in both populations indicates that additional loci must influence disease progression.

**Table 1. Frequency of *DLA-DRB1*/*-DQA1*/*-DQB1* haplotypes in collies and Shetland sheepdogs.**

COLLIE					
<i>DLA-DRB1</i> */ <i>-DQA1</i> */ <i>-DQB1</i> *	Cases (2n = 80)	Controls (2n = 370)	OR	95% CI	P-value
002:01/009:01/001:01	74	353	0.59	0.23–1.56	0.40
006:01/050:11/007:01	0	3	-	-	-
015:01/009:01/001:01	6	14	2.06	0.77–5.54	0.23
SHETLAND SHEEPDOG					
<i>DLA-DRB1</i> */ <i>-DQA1</i> */ <i>-DQB1</i> *	Cases (2n = 184)	Controls (2n = 410)	OR	95% CI	P-value
002:01/009:01/001:01 <sup>†</sup>	158	301	<b>2.20</b>	<b>1.38–3.52</b>	<b>0.0010</b>
015:01/009:01/001:01	1	6	0.37	0.04–3.08	0.4454
023:01/003:01/005:01	24	102	<b>0.45</b>	<b>0.28–0.73</b>	<b>0.0011</b>
Other haplotypes	1	1	-	-	-

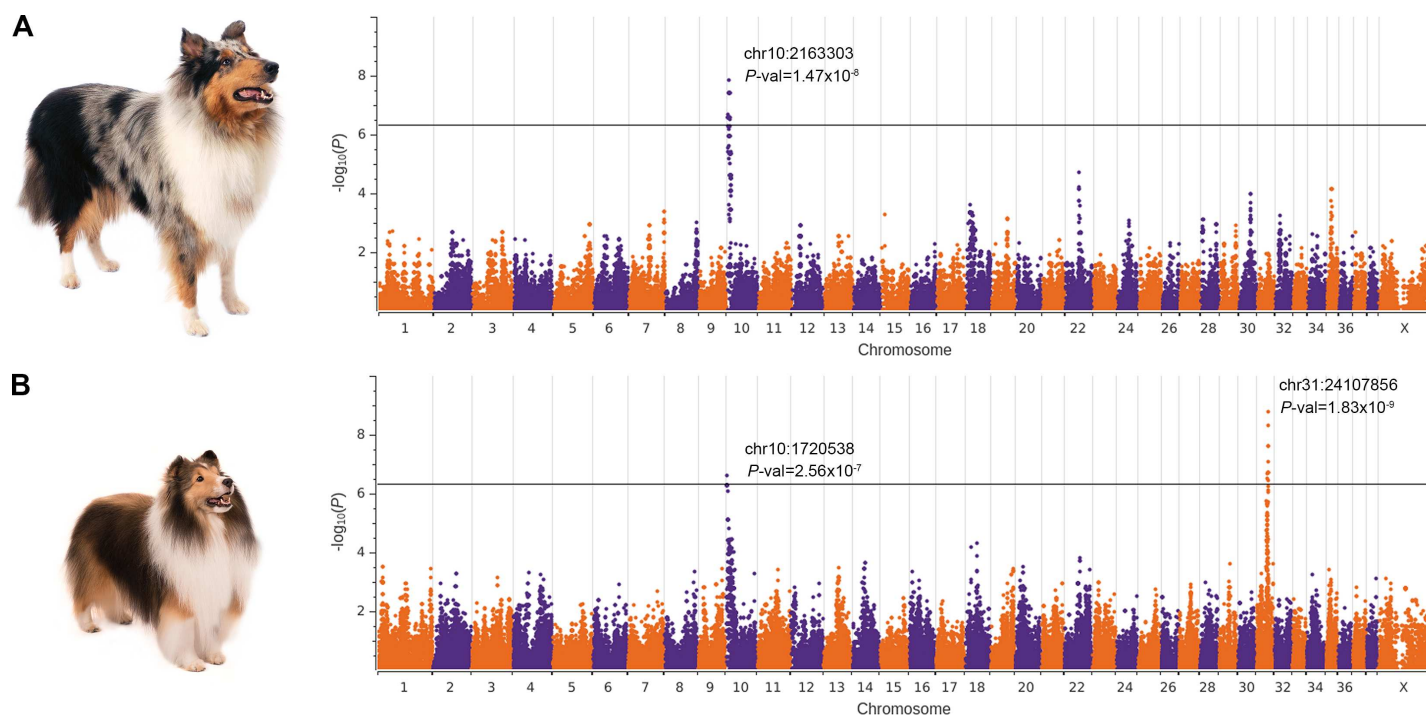
The number of times each haplotype was observed is reported for cases and controls with odds ratios (OR), 95% confidence intervals, and Fisher's exact two-tailed *P*-values. Significant statistics are bolded.

<sup>†</sup>71 out of 92 cases and 109 out of 205 controls were homozygous for the 002:01/009:01/001:01 haplotype (OR = 2.98, 95% CI = 1.70–5.21, *P*-val = 0.0001).

doi:10.1371/journal.pgen.1006604.t001

## Genome-wide association studies reveal signals on chromosomes 10 & 31

We conducted an independent GWAS for each breed, using a total of 97 cases (31 collies, 66 Shetland sheepdogs), 68 controls (23 collies, 45 Shetland sheepdogs), and 98,520 SNPs after filtering. In collies, a single signal (*P*-val =  $1.47 \times 10^{-8}$ ) composed of 17 SNPs at the centromeric end of chromosome 10 exceeded Bonferroni significance (Fig 1A). In Shetland sheepdogs, this



**Fig 1. Manhattan plots of DMS GWASs in two breeds.** (A) Collies: 31 cases and 23 controls. (B) Shetland sheepdogs: 66 cases and 45 controls. The  $-\log_{10}P$ -values (y-axis) for 98,520 Illumina SNPs are plotted against chromosome position (x-axis). The threshold for Bonferroni significance is shown as a black horizontal line. The *P*-value and position (canFam3) of the lead SNPs are reported.

doi:10.1371/journal.pgen.1006604.g001

association was replicated ( $P\text{-val} = 2.56 \times 10^{-7}$ ), and a second signal ( $P\text{-val} = 1.83 \times 10^{-9}$ ) composed of 11 SNPs surpassing Bonferroni significance was detected on chromosome 31 (Fig 1B). Both associations persisted in a combined breed analysis (chr10:  $P\text{-val} = 2.3 \times 10^{-12}$ , chr31:  $P\text{-val} = 3.95 \times 10^{-8}$ ) (S1 Table, S2 Fig); although the breeds possessed a common haplotype on chromosome 31, the Shetland sheepdogs appeared to drive this association. No associated SNPs were detected near the MHC loci on chromosome 12, likely a result of high homogeneity in our cohort and poor SNP coverage on the array [25].

On chromosome 10, 97% of all affected dogs were homozygous or heterozygous for the risk alleles of the lead SNPs. On chromosome 31, 88% of affected Shetland sheepdogs, but only 39% of affected collies, shared the risk alleles of the lead SNPs. As neither locus appeared to be independently necessary for disease development, we surveyed the extent of regional linkage disequilibrium (LD) to demarcate candidate intervals of ~1.33 Mb on chromosome 10 (Fig 2A) and ~696 kb on chromosome 31 (Fig 2B), harboring ~65 and 6 genes, respectively. The large size of the chromosome 10 region is attributed to lower recombination rates near the centromere and a dearth of informative SNPs.

### Identification of candidate variants in *PAN2* and *MAP3K7CL*

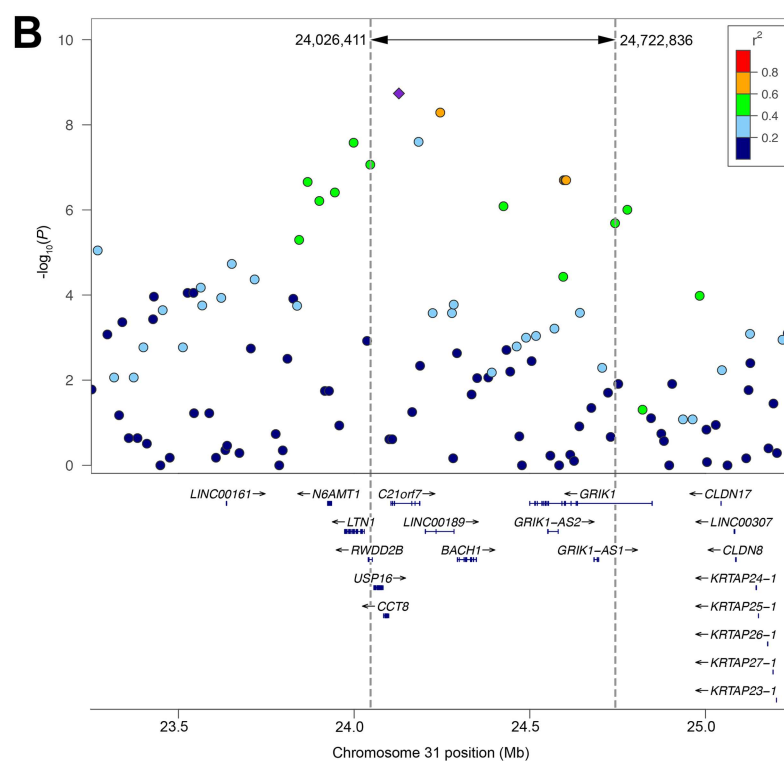
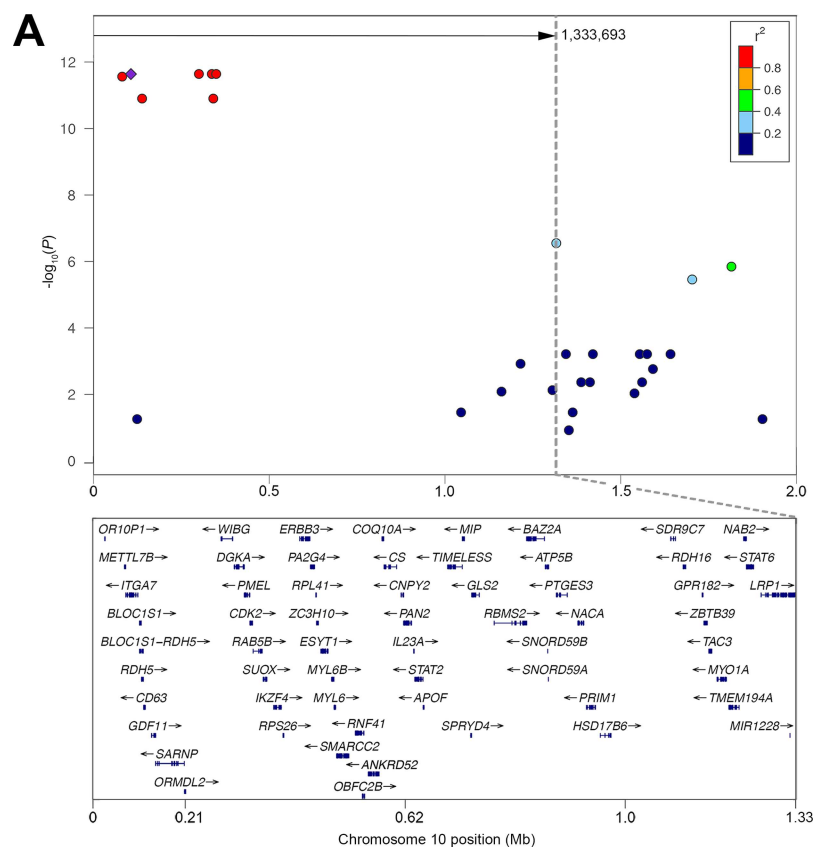
Whole genome resequencing was performed for four affected dogs (three collies and one Shetland sheepdog) and two unaffected collies, resulting in 17X to 41X coverage. Variants were filtered for those lying within our delineated intervals (chr10:1–1,333,693; chr31:24,026,411–24,722,836) and following the inheritance pattern of the most significantly associated SNPs in the affected dogs (S2 and S3 Tables). Five intergenic and two intronic variants were unique to these breeds (*i.e.*, not present in the Boxer reference genome, dbSNP, or 27 whole genome sequences from 16 other breeds); however, most were in repetitive regions and none were in conserved positions. Thus, the pathogenic variants were likely to be common polymorphisms, so we next prioritized variants within predicted exons and splice sites of genes expressed in skin for further study.

To confirm exon/intron boundaries predicted by Ensembl 79 and establish expression of variants in affected tissue, we generated RNAseq data. We obtained a minimum of 89 million reads per tissue from active skin lesions of two affected dogs (one collie and one Shetland sheepdog) and skin from one unaffected Australian shepherd dog. All genes expressed in skin within the candidate regions were also expressed in affected tissues. Seventeen exonic variants were expressed; seven of these were nonsynonymous and evaluated using *in silico* programs [31–33] (Table 2).

We genotyped a subset of our mapping population for three nonsynonymous SNPs on chromosome 10 (*ANKRD52* g.565958G>C, *PAN2* g.627760G>A, and *STAT6* g.1239562G>A) that were predicted to be deleterious or probably damaging by more than one *in silico* program. The *ANKRD52* and *PAN2* variants were more strongly associated with DMS than the lead SNP. These variants were in perfect linkage disequilibrium with each other; however, *PAN2* g.627760G>A was assigned damaging scores with higher confidence by *in silico* programs (Table 2). We therefore focused further studies on *PAN2* g.627760G>A, encoding p.Arg492Cys (XP\_531635.3), although *ANKRD52* cannot be excluded. On chromosome 31, we genotyped Shetland sheepdogs for a SNP (g.24132343A>C) and an indel (c.383\_392ACTCCACAAA>GACT, XM\_846337.4), both located in a 5' non-coding exon of *MAP3K7CL*. Only the indel was associated with DMS (Table 2, S3 Fig). In an expanded, combined population (132 affected and 390 unaffected collies and Shetland sheepdogs), both the *PAN2* ( $P\text{-val} = 2.08 \times 10^{-35}$ ) and *MAP3K7CL* ( $P\text{-val} = 1.45 \times 10^{-33}$ ) variants were highly associated with DMS (S4 Table).

*PAN2* (or *USP52*) encodes the catalytic subunit of the poly(A) nuclease deadenylation complex (PAN2-PAN3) and is one of two exonucleases involved in mRNA degradation in eukaryotes





**Fig 2. Regional plots depicting LD at two associated loci.** Association results ( $-\log_{10}P$ ) for (A) combined breed GWAS on chromosome 10 and (B) Shetland sheepdog GWAS on chromosome 31 are color-coded based on pairwise LD ( $r^2$ ) with the lead SNPs (purple diamond). Dotted lines demarcate the candidate interval. Genes within the region (based on hg18) are shown below the plot with chromosome positions reported in canFam3. On chromosome 10, the gene track is zoomed in to represent the candidate interval only; 4 genes were omitted from the figure (*DNAJC14*, *LOC440104*, *MMP19*, and *SLC39A5*).

doi:10.1371/journal.pgen.1006604.g002

[34,35]. Deadenylation plays a role in translational regulation of inflammatory response [36]. Independent of this function, PAN2 also stabilizes *HIF1A* transcripts via their 3'-UTR, which contain AU-rich elements (AREs), and may be involved in regulating other transcripts having AREs [37]. *HIF1A*, a key regulator of inflammation [38], and other ARE-containing transcripts, such as *IL-6* [39], are upregulated in JDM [40, 41]. *PAN2* is widely expressed and highly evolutionarily conserved [35]; human (NP\_001120932.1) and dog (XP\_013972628.1) amino acid sequences share 98% identity.

*MAP3K7CL* (also known as *TAK1L* or *C21orf7*) is a poorly studied kinase gene that is transcriptionally active in immunological tissues and expressed primarily in peripheral blood leukocytes [42,43]. Human (NP\_001273546.1) and dog (XP\_013965340.1) *MAP3K7CL* protein sequences share >90% identity. The transcription factors RUNX3 and EP300 bind the 5' non-coding exon of human *MAP3K7CL* (UCSC Genome Browser ENCODE Transcription Factor ChIP-seq track). In this exon, the c.383\_392ACTCCACAAA>GACT indel causes the loss of six conserved base pairs, omitting a RUNX3 binding motif ( $P$ -val =  $2.07 \times 10^{-3}$  from TOMTOM [44]) (S3 Fig). RUNX3 has known roles in inflammatory response (e.g., thymopoiesis [45,46]

**Table 2. Expressed exonic changes on chromosomes 10 and 31.**

Position	Gene	A1/A2	AA change	Poly-Phen2	PANTHER	SIFT	P-value
<b>CHROMOSOME 10</b>							<b>Cases n = 93 Controls n = 63</b>
95042	<i>ITGA7</i>	T/C	Val/Val	-	-	-	-
121835	<i>RDH5</i>	G/A	Ala/Ala	-	-	-	-
222809	<i>MMP19</i>	T/G	Gln/Pro	0	91	0.43	-
331135	<i>SUOX</i>	A/G	His/Arg	0	85	N/A	-
<b>565958</b>	<b><i>ANKRD52</i></b>	<b>G/C</b>	<b>Ser/Cys</b>	<b>0.92</b>	<b>361</b>	<b>0.02</b>	<b><math>1.93 \times 10^{-17}</math></b>
<b>627760</b>	<b><i>PAN2</i></b>	<b>G/A</b>	<b>Arg/Cys</b>	<b>1</b>	<b>1628</b>	<b>0</b>	<b><math>1.93 \times 10^{-17}</math></b>
1127082	<i>RDH16</i>	C/T	Ala/Ala	-	-	-	-
1228277	<i>NAB2</i>	G/C	Gly/Arg	0	85	N/A	-
1239562	<i>STAT6</i>	G/A	Thr/Met	0.98	361	0.05	$1.52 \times 10^{-4}$
1286150	<i>LRP1</i>	C/T	Asn/Asn	-	-	-	-
1333693	<i>LRP1</i>	G/A	Gln/Gln	-	-	-	-
<b>CHROMOSOME 31</b>							<b>Cases n = 63 Controls n = 45</b>
24068039	<i>CCT8</i>	T/C	Ala/Ala	-	-	-	-
<b>24132273</b>	<b><i>MAP3K7CL</i></b>	<b>Indel/-</b>	<b>5' UTR</b>	<b>N/A</b>	<b>N/A</b>	<b>N/A</b>	<b><math>2.09 \times 10^{-11}</math></b>
24132343	<i>MAP3K7CL</i>	A/C	5' UTR	N/A	N/A	N/A	1
24292521	<i>BACH1</i>	A/G	Asn/Asp	0.007	324	0.21	-
24294659	<i>BACH1</i>	A/G	Glu/Glu	-	-	-	-
24295208	<i>BACH1</i>	A/G	Pro/Pro	-	-	-	-

Alleles 1 (A1 = minor allele) and 2 (A2 = reference allele) are reported. Amino acid changes are based on the dog reference genome.  $P$ -values are reported for matched populations. Variants with PolyPhen2 scores ranging from 0.85–1, PANTHER preservation time (in millions of years) >450my, and SIFT scores ranging from 0–0.05 are considered deleterious/probably damaging. Variants more strongly associated with DMS than the most associated SNPs from the array are bolded.

doi:10.1371/journal.pgen.1006604.t002



and the TGF- $\beta$  signaling cascade [47]), and it has been directly implicated in a number of immune-related diseases [48–50]. Furthermore, SNPs disrupting RUNX binding motifs in target genes confer susceptibility to autoimmune rheumatic diseases, including psoriasis [51] and systemic lupus erythematosus [52].

### Analysis of three-locus genotypes reveal gene-gene interactions

We next considered three-locus genotypes in our expanded, combined population (132 affected and 390 unaffected dogs) where *A* = the *PAN2* variant encoding p.Arg492Cys, *B* = *MAP3K7CL* c.383\_392ACTCCACAAA>GACT, and *C* = *DLA-DRB1*\*002:01, lower-case letters denote wild type alleles (*c* represents any alternate allele of *DLA-DRB1*). Only 4% of dogs possessed a three-locus genotype with *cc*, barring further analysis of those nine genotypes. We considered nine of the remaining genotypes to be low-risk, as less than 6% of dogs with these allelic combinations had DMS (Table 3). Among healthy dogs (Fig 3A), the most frequently observed genotypes

**Table 3. Distribution of three-locus genotypes in 132 cases and 390 controls with penetrance, 95% confidence intervals, and *P*-values.**

Genotype	Collies		Shetland sheepdogs		Combined			
	Cases (n = 40)	Controls (n = 185)	Cases (n = 92)	Controls (n = 205)	Penetrance (%)	Risk*	95% Confidence Interval	<i>P</i> -value**
<i>aabbcc</i>	0	1	0	1	-	-	0–0.842	1
<i>aabbCc</i>	0	5	0	18	-	low	0–0.148	<b>0.0023</b>
<i>aabbCC</i>	1	43	1	17	-	low	0.004–0.112	<b>0</b>
<i>Aabbcc</i>	0	0	0	3	-	-	0–0.708	0.577
<i>AabbCc</i>	0	3	1	15	-	low	0.001–0.260	0.060
<i>AabbCC</i>	2	71	2	23	-	low	0.011–0.101	<b>0</b>
<i>aaBbcc</i>	0	0	0	5	-	low	0–0.522	0.399
<i>aaBbCc</i>	0	1	0	21	-	low	0–0.154	<b>0.002</b>
<i>aaBbCC</i>	0	10	0	32	-	low	0.084	<b>0</b>
<i>AaBbcc</i>	0	0	0	1	-	-	0–0.975	1
<i>AaBbCc</i>	0	0	0	14	-	low	0–0.232	<b>0.028</b>
<i>AaBbCC</i>	0	13	2	23	-	low	0.006–0.177	<b>0.002</b>
<i>AAbbcc</i>	1	0	0	0	-	-	0.025–1	0.253
<i>AAbbCc</i>	4	6	1	1	42	moderate	0.152–0.723	0.194
<i>AAbbCC</i>	15	31	7	3	39	moderate	0.265–0.532	<b>0.021</b>
<i>aaBBcc</i>	0	0	0	1	-	-	0–0.975	1
<i>aaBBCc</i>	0	0	0	6	-	low	0.0459	0.348
<i>aaBBCC</i>	0	0	6	7	46	moderate	0.192–0.749	0.107
<i>AaBBcc</i>	0	0	2	1	-	-	0.094–0.992	0.160
<i>AaBBCc</i>	0	0	5	5	50	moderate	0.187–0.813	0.136
<i>AaBBCC</i>	3	0	23	3	90	high	0.762–0.978	<b>0</b>
<i>AABbcc</i>	0	0	3	0	-	-	0.292–1	<b>0.016</b>
<i>AABbCc</i>	0	0	2	4	33	moderate	0.043–0.777	0.647
<i>AABbCC</i>	11	1	11	1	92	high	0.730–0.990	<b>0</b>
<i>AABBcc</i>	0	0	0	0	-	-	-	-
<i>AABBCc</i>	0	0	7	0	100	high	0.590–1	<b>0.0001</b>
<i>AABGCC</i>	3	0	19	0	100	high	0.846–1	<b>0</b>

Significant values in bold.

\*Risk interpretations were only made for three-locus genotypes observed at least five times.

\*\**P*-values are calculated for each genotype compared to the population as a whole (0.253, 132 affected dogs out of 522 total dogs).

doi:10.1371/journal.pgen.1006604.t003

were *AabbCC* (24%) and *aabbCC* (15%). Based on penetrance, we classified five genotypes as conferring moderate risk (33–50%) and four as high risk (90–100%) for DMS. All cases possessed at least two risk alleles and all but one were homozygous for at least one risk allele. The most common genotypes of affected dogs (Fig 3B) were *AaBBCC* (20% of cases), followed by *AAbbCC*, *AABbCC*, and *AABBCC* (17% of cases each).

Interestingly, only affected dogs possessed *AABBCc* or *AABBCC* ( $n = 29$ ), indicating that DMS is fully penetrant in dogs having these combinations. Epistasis plots illustrated that genotypes with at least one *a* or *b* allele confer a lower probability of disease when a *c* allele is present, compared to their *CC* counterparts (Fig 3, compare 3C and 3D). The plots also depicted a greater probability of disease than expected under a strictly additive model, providing evidence for additive-by-additive epistasis between the chromosome 10 and 31 loci [53,54]. We noted at least one ARE in *MAP3K7CL*, presenting a mechanism for interaction with *PAN2*. No difference in gene interactions was observed between the sexes (S4 Fig).

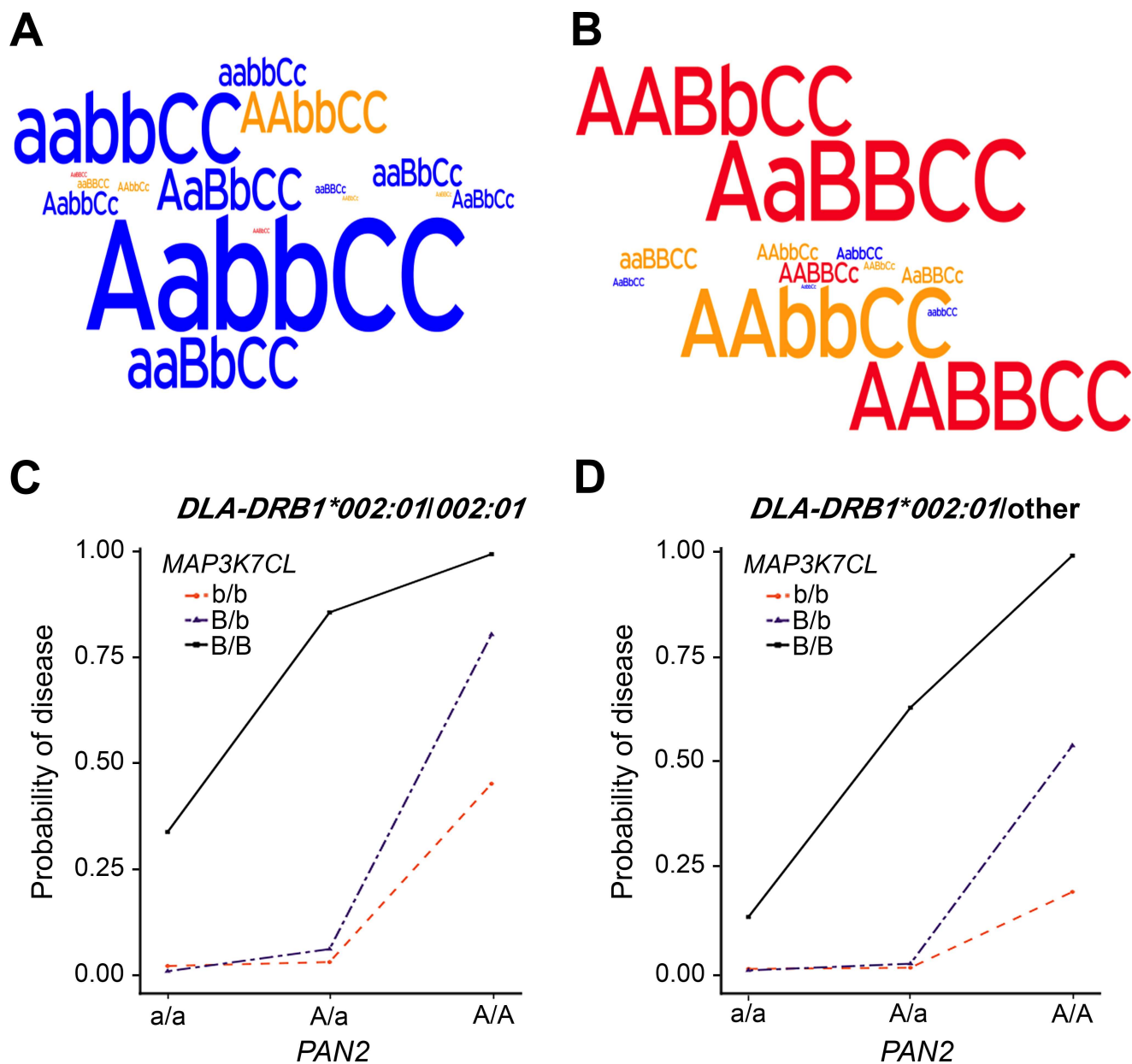
Information regarding age at onset or diagnosis was available for 91 dogs. We compared dogs having two, three, or four risk alleles across *PAN2* and *MAP3K7CL* and observed an inverse correlation between age of onset and number of risk alleles (S5 Fig). Dogs having four risk alleles developed DMS at a significantly younger median age (5 months) than did dogs with only two risk alleles (18.5 months). The complete penetrance of *AABB* genotypes, combined with an early age of onset, suggest that these dogs may be hypersensitive to common-place environmental stimuli (e.g., routine puppy vaccinations).

## Collies and Shetland sheepdogs have uniquely high frequencies of associated alleles

All three identified variants associated with DMS are polymorphisms present in several breeds, raising the question: why are other breeds rarely, if ever, affected by DMS? We genotyped five or more unrelated individuals from each of 30 diverse breeds for all three loci (Fig 4). The only other breeds to possess all three risk alleles were Cardigan Welsh corgis and Cairn terriers. Three Jack Russell terriers had moderate-risk genotypes (*AAbbCc*), as did one Cardigan Welsh corgi (*AAbbCc*); both breeds are occasionally diagnosed with dermatomyositis-like disease [55,56]. None of the 229 individuals possessed a high-risk genotype (S5 Table). Interestingly, Labrador retrievers had both *B* and *C*, which could have enabled moderate or high risk genotypes (*AaBBCc* or *AaBBCC*) in puppies from the outcross mating described by Haupt et al. [14].

Combined frequencies of risk alleles in other breeds were dramatically lower than those observed among collie and Shetland sheepdog populations, and homozygosity for a risk allele (a characteristic of all moderate- to high-risk genotypes) was rare. Additionally, breeds having a high frequency of one risk allele had few or no risk alleles at the other loci. For example, Cairn terriers had a high frequency of *A* (75%) but low frequencies of *B* (19%) and *C* (3%), and no high- or moderate-risk genotypes were observed among 18 individuals. These findings suggest that independently the polymorphisms are neither deleterious nor selected against.

It is likely that recent artificial selection for phenotypes shared by collies and Shetland sheepdogs led to increased frequencies of *A*. We propose that persistence of *A* in these two breeds is attributed to linkage disequilibrium ( $D' = 0.998$ ) between wildtype *PAN2* (*a*) and another chromosome 10 allele, *Merle* of *PMEL*. In heterozygosity, *Merle* causes a popular pigmentation pattern (see collie in Fig 1A), but homozygosity for the allele results in severe hypopigmentation often with auditory and ocular defects [58]. Wildtype *PMEL* occurred on chromosomes with either *A* or *a*, whereas *Merle* was only found in conjunction with *a*. Accordingly, the *Merle* phenotype was underrepresented in affected dogs ( $P$ -value = 0.0018),

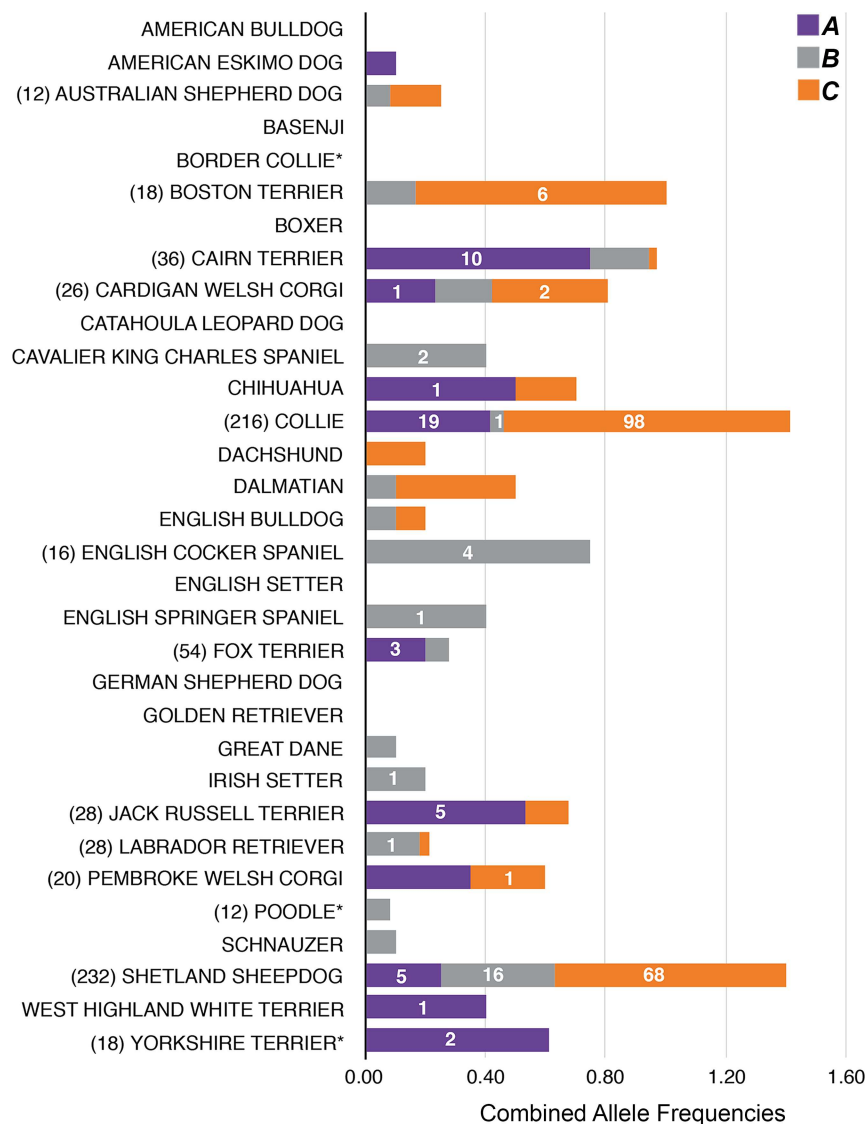


**Fig 3. Three-locus genotype disease probabilities and frequencies (132 cases, 390 controls).** Word clouds of three-locus genotypes illustrate frequency differences between (A) controls and (B) cases with high-risk genotypes in red, moderate-risk in yellow, and low-risk in blue. Probability of disease (y-axis) for all combinations of PAN2 (x-axis) and MAP3K7CL genotypes are plotted in dogs (C) homozygous and (D) heterozygous for *DLA-DRB1\*002:01*.

doi:10.1371/journal.pgen.1006604.g003

64% of which were homozygous for A. Consistent selection for heterozygosity (but not homozygosity) for *Merle* would simultaneously encourage maintenance of both *PAN2* alleles. To our knowledge, there are no loci on chromosome 31 under positive selection for maintenance of a characteristic phenotype of collies and/or Shetland sheepdogs.

Across five collie genomes, we observed a 1.2 Mb selective sweep on chromosome 12 (S6 Fig) encompassing the MHC class II loci and leading to near fixation of the 002:01/009:01/001:01 haplotype. We suggest that essentially all purebred collies have increased susceptibility for DMS, conferred by homozygosity for allele C. The Shetland sheepdog population has



**Fig 4. Frequencies of DMS-associated polymorphisms across breeds.** The number of alleles used to calculate the frequencies is shown to the left of the breed name, if > 10. Risk allele frequencies (x-axis) for each gene are represented by a single, segmented colored bar: the *PAN2* variant encoding p.Arg492Cys is represented in purple, *MAP3K7CL* c.383\_392ACTCCACAAA>GACT in gray, and *DLA-DRB1\*002:01* in orange. The number of homozygous dogs observed is reported on the corresponding color bar. Asterisks denote breeds lacking *DLA-DRB1\*002:01* in this study, but which were previously reported to possess the allele [57].

doi:10.1371/journal.pgen.1006604.g004

retained a less common second haplotype that permits heterozygosity at the DLA loci, associated with a lower risk for developing DMS. Ironically, this reduced risk may have masked the presence of otherwise high-risk genotypes (*i.e.*, *AABb* and *AaBB*) and hindered selection against *A* and *B* alleles.

We observed striking allele frequency differences between the two affected breeds at *PAN2* and *MAP3K7CL*: collies had a higher frequency of *A*, 42% (25% in Shetland sheepdogs), whereas Shetland sheepdogs had a higher frequency of *B*, 38% (5% in collies) (Fig 4). Consequently, the frequency of observed allelic combinations varied between the breeds. The most common genotypes in healthy dogs were *aaBbCC* in Shetland sheepdogs, 16% (5% in collies),

and *AabbCC* in collies, 38% (11% in Shetland sheepdogs). Among affected dogs, *AaBBCC* predominated in Shetland sheepdogs, 25% (8% in collies), whereas *AabbCC* was the most frequent genotype in diseased collies, 38% (8% in Shetland sheepdogs). The latter finding is interesting given that *AabbCC* is only a moderate-risk genotype. Among collies having this genotype, 67% were unaffected by age 8, whereas only 30% of Shetland sheepdogs with this genotype were disease-free. This discrepancy in disease probabilities between breeds was unique to this genotype. Further studies will be necessary to determine if other loci confer additional risk for or protection from DMS.

The contribution of alleles from multiple loci explains the spontaneous appearance of the disease in lines with no prior history [14] and has hindered elimination of DMS in the absence of a genetic test. For example, a mating between two healthy dogs having low risk genotypes (*e.g.*, *AaBbCC* x *AaBbCC*) can produce puppies with low, moderate, or high risk for DMS. This study has led to the first three-locus genetic test for a complex disease of dogs, which will allow breeders to carefully reduce the frequency of *A* and *B* among collie and Shetland sheepdog populations while preserving genetic diversity and desirable breed characteristics.

In a canine model of JDM, we have identified a complex pattern of causation involving three independent loci, two of which offer new targets for exploration in JDM. Furthermore, these data provide support for the involvement of genetic risk factors independent of the MHC in human inflammatory myopathies. While further experiments are necessary to determine the exact contribution of the chromosome 10 and 31 loci, our findings suggest that DMS may result from an inability to properly regulate inflammatory response. This work highlights the utility of the canine model for unraveling genetic susceptibility conferred by common polymorphisms and/or gene-gene interactions in complex diseases.

## Methods

### Ethics statement

All samples were obtained with informed consent according to protocols approved by the Clemson University Institutional Review Board (IBC2008-17) and IACUC (2012–039).

**Study population.** Three populations were assembled: DMS-affected dogs (92 Shetland sheepdogs, 40 collies), control dogs for GWAS (45 Shetland sheepdogs, 23 collies), and unaffected dogs (160 Shetland sheepdogs, 162 collies). Affected dogs were diagnosed either through histopathology (75 Shetland sheepdogs, 30 collies) or by a veterinarian based on clinical signs and elimination of demodectic mange, a differential diagnosis. Pedigrees were collected when available; some samples were obtained from affected dogs surrendered to rescue organizations without paperwork. Twenty percent of affected dogs were collected internationally, and in each case a regionally-matched control sample was obtained.

Control dogs for GWAS were eight years of age or older at time of collection, had no known family history of DMS, had no personal history of skin disease, and were unrelated to other study participants within two (most often three) generations. Pedigrees were obtained from all control dogs. The population of unaffected dogs had no clinical signs of DMS and were eight years of age or older at time of collection. This subset was collected without regard to family history of DMS, presence of other skin disorders, or relationship to other study participants. Archival samples from 229 dogs of 30 other breeds were not closely related to each other to our knowledge.

Whole blood or buccal cells were obtained from each dog, and genomic DNA was isolated according to the Puregene DNA Isolation protocol (Gentra). Skin punch biopsies from active lesions or healthy tissue were also obtained from one Shetland sheepdog, one collie, and one Australian shepherd dog and preserved in RNAlater (Ambion).

**DLA class II genotyping.** The hypervariable regions of *DLA-DRB1*, *-DQA1* and *-DQB1* were sequenced and genotyped according to protocols previously described [59–61]. *DLA-DQA1* was largely uninformative in collies and was inferred. Association of a haplotype or allele state with DMS was assessed through Fisher's exact tests, using VassarStats (Web Site for Statistical Computation, Vassar College, Poughkeepsie, NY).

**Genome-wide association and LD analyses.** Genotyping was performed for 166 dogs (85 males and 81 females) using the Illumina CanineHD BeadChip, containing 173,662 SNPs. One sample having call rates <95% was excluded. SNPs having call rates <95%, minor allele frequencies <5%, and/or significant deviation from Hardy-Weinberg equilibrium ( $P$ -value <0.0001) in the control dogs were excluded. No evidence of genomic inflation was observed in the combined Shetland sheepdog and collie analysis ( $\lambda = 1.04$ ). Fisher's exact  $P$ -values were calculated under a dominant model.

LD pairwise analysis was performed to calculate  $r^2$  values, which were plotted using LocusZoom [62]. We calculated  $r^2$  values between SNPs on chromosome 10 using all controls. For chromosome 31,  $r^2$  values were calculated using only control Shetland sheepdogs because no association was detected in collies alone. Assuming that the pathogenic variant would be in high LD with the lead SNP, candidate regions were characterized by SNPs with pairwise  $r^2 \geq 0.6$  and defined by the first flanking associated SNPs in lower LD. Filtering and statistical analyses were conducted with SNP & Variation Suite v8 (SVS, Golden Helix). All chromosome positions throughout the text are reported in CanFam3.1.

**Whole genome resequencing.** Three affected and two control collies were selected for whole genome resequencing along with one affected Shetland sheepdog. Genomic DNA fragments of approximately 500bp were gel size selected for each sample and sequenced on two lanes of an Illumina HiSeq 2000, generating 2x100 (collies) and 2x125 bp (Shetland sheepdog) paired-end reads. 531 to 997 million total reads were generated for the five collies. Over one billion total reads were generated for the Shetland sheepdog. Paired reads were aligned to the indexed reference (CanFam3.1) with Bowtie2 [63] under sensitive parameters. The alignments were sorted and indexed with SAMtools [64,65] to be visualized in the Interactive Genomics Viewer [66].

**RNA isolation and sequencing.** Total RNA was extracted from 30–40mg of skin punch biopsy tissue from active lesions (collie and Shetland sheepdog) or healthy skin (Australian shepherd) using the ToTALLY RNA kit (Ambion), according to the manufacturer's protocols. RNA samples were treated with DNase to remove DNA contamination, using the DNA-free kit (Ambion). Samples were quantitated on a NanoDrop 1000 spectrophotometer (Fisher Scientific).

Three RNAseq libraries were constructed per dog using normalized total RNA and the manufacturer's protocol for one of the following: TruSeq RNA library prep kit v2.0 (Illumina) or TruSeq stranded total RNA library prep kit (Illumina). An Agilent Bioanalyzer 2100 was used for size validation. Each sample was sequenced at 2x125 bp paired-end on an Illumina HiSeq 2500 to a depth of at least 22 million reads.

FastQC from the Babraham Institute was used to assess read quality before and after pre-processing by Trimmomatic [67], which removed low quality bases and remaining sequence adapters. Trimmed reads for each sample were aligned to CanFam3.1 using gsnap [68], and SAMtools [64,65] was used to generate sorted and indexed bam files.

**Variant filtering and genotyping.** SAMtools and BCFtools [64,65] were used to generate variant call files for each sample, which were analyzed in SVS. On chromosome 10 (1–1,333,693 bp) variants that were heterozygous in affected collies 1 & 2 and homozygous in affected collie 3 and the affected Shetland sheepdog (*i.e.*, reference/reference or alternate/alternate), consistent with the allele states at the lead SNPs, were selected. On chromosome 31



(24,026,411–24,722,836 bp), all homozygous variants in the Shetland sheepdog were selected. Ensembl 79 was used to identify variants lying within predicted exons and 10 bp flanking sequences to capture splice sites. RNAseq data were manually inspected in IGV to determine whether predicted exonic variants were expressed in the affected dogs. Alternate variants were investigated using dbSNP and 27 genomes of 16 other breeds (either sequenced as part of ongoing studies or shared by other research groups) to determine whether any were unique to collies and Shetland sheepdogs. SIFT [31], PolyPhen [32], and PANTHER [33] were used to predict the impact of the amino acid substitutions. Genotyping of variants in additional dogs was accomplished by restriction digest assays or Sanger sequencing (S6 Table).

**Age of onset.** Because age is not distributed as a normal random variable, we made use of the Weibull distribution, a density commonly used for time-to-event data [69]. Our goal was to estimate the mean and median age of onset, along with their confidence intervals, to facilitate comparisons across the three genotypic groups.

Taking advantage of a Bayesian strategy, we assume  $y_{ij} \sim \text{Weibull}(r, e^{-b_i})$ , where  $y_{ij}$  is the observed age at diagnosis for the  $j$ -th ( $j = 1, 2, \dots, n_i$ ) dog in the  $i$ -th ( $i = 1, 2, \dots, 3$ ) genotypic class and  $b_i$  is the effect of the  $i$ -th genotypic class. With this representation of the scale ( $r$ ) and shape ( $e^{-b_i}$ ) parameters of the Weibull density, the median of the  $i$ -th genotypic class is  $e^{-b_i} \log(2)^{1/r}$  and the mean of the  $i$ -th genotypic class is  $e^{-b_i} \Gamma(1 + 1/r)$ .

In addition, we assumed a prior density for the unknown genetic parameters ( $b_i$ ) to have a null mean and common variance  $\sigma$  (i.e.,  $(b_i \sim N(0, \sigma))$ ). We consider the weakly informative prior for  $\sigma \sim \text{Cauchy}(0, 25)$  [70]. For the scale parameter,  $r$ , we assume the weakly informative prior of  $r \sim \text{exponential}(0.001)$ . The Monte Carlo Markov Chain (MCMC) sampling process was run in 4 chains through the public-domain software Stan [71], with each chain based on 50,000 total samples, and the first 20,000 were removed as part of the warm-up process, then thinned to every 20th sample, resulting in a MCMC sample of 6,000 values [71]. Convergence to the posterior density was evaluated by the Gelman-Rubin test statistic, where values less than 1.05 indicate that the MCMC sampling process was adequate to the data evaluated [72].

**Genetic interaction analyses.** All affected ( $n = 132$ ) and healthy ( $n = 390$ ) dogs were used to investigate possible interactions between the *PAN2* variant encoding p.Arg492Cys and *MAP3K7CL* c.383\_392ACTCCACAAA>GACT in dogs homozygous for 002:01/009:01/001:01 vs. dogs heterozygous for the haplotype. Three-locus genotypes observed in fewer than five dogs were excluded. For a given disease state (cases vs. controls) and a given pair of genotypic classes  $i$  ( $i = A/A, A/a, \text{ or } a/a$ ) and  $j$  ( $j = B/B, B/b, \text{ or } b/b$ ),  $n_{\text{cases } ij} \sim \text{Binomial}(n_{\text{cases } ij} + n_{\text{controls } ij}, p_{ij})$  where  $n_{\text{cases } ij}$  is the number of observed cases in the combination of the genotypic class of locus  $i$  and that of locus  $j$ ,  $n_{\text{controls } ij}$  is the number of unaffected dogs in the combination of genotypic class of locus  $i$  and that of locus  $j$ . Finally,  $p_{ij}$  is the probability of disease for the combined genotypic classes of  $i$  and  $j$ .

Estimation of the unknown elements of our model must ensure estimates of  $p_{ij}$  within the interval  $[0, 1]$ , recognizing that several genotypic classes have zero or few cases. We utilized a hierarchical Bayesian framework facilitated through the public-domain software Stan [71] to address this problem [73]. Stan can be accessed through the public-domain language R [74]. Log-odds was used to estimate  $p_{ij}$ , i.e.,  $\log(p_{ij}/(1 - p_{ij})) = \text{intercept} + \text{add}A_i + \text{dom}A_i + \text{add}B_j + \text{dom}B_j + \text{add} \times \text{add}_{ij} + \text{add} \times \text{dom}_{ij} + \text{dom} \times \text{add}_{ij} + \text{dom} \times \text{dom}_{ij}$ , where intercept represents a term common to all genotypic classes,  $\text{add}A_i$ ,  $\text{add}B_j$ ,  $\text{dom}A_i$  and  $\text{dom}B_j$  represent the additive and dominance terms respectively for loci  $A$  and  $B$ , and  $\text{add} \times \text{add}_{ij} + \text{add} \times \text{dom}_{ij} + \text{dom} \times \text{add}_{ij} + \text{dom} \times \text{dom}_{ij}$  represent the four possible epistatic interaction terms for all possible additive and dominance combinations [53].

We assumed a prior density for the intercept and unknown genetic parameters, with null mean and common variance  $\sigma$  (*i.e.*,  $N(0, \sigma)$ ). Subsequently, we consider the weakly informative prior for  $\sigma \sim \text{Cauchy}(0, 25)$  [70]. We used the same MCMC parameters here as described above.

**Chromosome 12 selective sweeps.** We identified all SNPs on chromosome 12 present in five collie genomes (3 cases and 2 controls) and used a creeping window size  $\leq 1$  Mb to identify runs of homozygosity [75]. Windows containing fewer than 50 SNPs were excluded and gaps  $> 10$  kb between SNPs were ignored. The heterozygosity (Hp) statistic was calculated for all windows and Z-transformed, making the average Hp value equal to zero and the standard deviation equal to 1. The  $-ZHp$  distribution was plotted in R to show putatively selected regions greater than 3.4 standard deviations from the mean.

**PMEL.**  $D'$  between *PMEL* and *PAN2* was calculated as a measure of LD using all dogs for which coat pattern phenotypes were available (112 cases, 385 controls). A total of 98 dogs (10 cases and 88 controls) were described as merle-patterned, a semi-dominant trait caused by the *Merle* (*M*) allele of *PMEL*, and assumed to possess the *Mm* genotype. Computation of  $D'$ , along with a test of significance from zero, was facilitated through the package *genetics* [76], a public domain program that is part of the R language [74].

## Accession numbers

SNP data are available in dbSNP under BioProject number PRJNA338128. All whole genome and transcriptome data generated for this study were deposited in SRA genomes under accession number SRP081080. Accession numbers for eight of the 27 other breeds used in variant filtering are SRX1360633, SRX1360635, SRX1360637, SRX1360639, SRX1022256, SRX1022262, SRX1022286, and SRP081080.

## Supporting information

**S1 Fig. Clinical presentation of dermatomyositis.** Canine dermatomyositis is a vasculopathy that initially manifests as cutaneous lesions across the bony prominences of the face, tail tip, limbs, and feet, shown here. Some dogs develop alopecia and more extensive lesions over time, resulting in dermal scarring associated with erythema and mottled pigmentation. (PDF)

**S2 Fig. Manhattan and quantile-quantile plots of combined DMS GWAS (97 cases vs. 68 controls).** The  $-\log_{10}P$ -values (y-axis) for 98,520 Illumina SNPs are plotted against chromosome position (x-axis). The threshold for Bonferroni significance is shown as a black horizontal line. The  $P$ -value and position (canFam3) of the lead SNPs are reported. The Q-Q plot is boxed in purple and plots observed vs. expected Fisher's exact  $-\log_{10}P$ -values. The inflation factor ( $\lambda$ ) is shown on the Q-Q plot. (PDF)

**S3 Fig. *MAP3K7CL* indel with conservation and *RUNX3* binding motif.** UCSC 100 Vertebrates track for human chr21:29,130,846–29,130,860 showing the G insertion and seven base pair deletion created by *MAP3K7CL* indel (Dog c.383\_392ACTCCACAAA>GACT). Bases in gray differ from the dog reference sequence. The canine sequence is highlighted in yellow. The *RUNX3* binding motif is underlined. (PDF)

**S4 Fig. Three-locus genotype disease probabilities by sex.** Probability of disease (y-axis) for all combinations of *PAN2* (x-axis) and *MAP3K7CL* genotypes are plotted in dogs (73 affected and 145 unaffected males; 59 affected and 245 unaffected females) (top) homozygous and

(bottom) heterozygous for *DLA-DRB1\*002:01*.  
(PDF)

**S5 Fig. Median age of onset for combinations of *PAN2* and *MAP3K7CL* genotypes.** Median age of onset is plotted for genotypes consisting of 2 (AAbb, aaBB, AaBb), 3 (AABb, AaBB), and 4 (AABB) risk alleles at *PAN2* and *MAP3K7CL*. CC and Cc genotypes were combined for analyses. Number of individuals is shown to right.  
(PDF)

**S6 Fig. Collie selective sweeps on chromosome 12.**  $ZH(p)$  values for all creeping windows containing 50 or more SNPs are plotted against chromosome position. Creeping windows are  $\leq 1$  Mb.  
(PDF)

**S1 Table. SNPs exceeding Bonferroni significance ( $5.08 \times 10^{-7}$ ).**  
(PDF)

**S2 Table. Chromosome 10 variants segregating with the lead SNPs in the affected dogs.**  
(PDF)

**S3 Table. Chromosome 31 variants segregating with the lead SNPs in the affected Shetland sheepdog.**  
(PDF)

**S4 Table. Frequency of *PAN2* and *MAP3K7CL* genotypes.**  
(PDF)

**S5 Table. Three-locus genotypes for 229 individuals of 30 breeds.**  
(PDF)

**S6 Table. Primers and genotyping method for variants.**  
(PDF)

## Acknowledgments

The authors wish to thank Dr. Christine Rees and Sherry Lindsey for contributing samples, as well as the many dog owners and veterinarians. We thank Drs. Gary Johnson and Kathryn Meurs for providing genomes from unaffected dogs. We thank Drs. Mike Vaughan, Michael Sehorn, and Stephen Kresovich for critical reading of the manuscript.

## Author contributions

**Conceptualization:** LAC KLT JME.

**Data curation:** REN KLT JME.

**Formal analysis:** JME TRF ANSM.

**Funding acquisition:** LAC KLT.

**Investigation:** JME REN CMH KJA.

**Project administration:** LAC.

**Supervision:** JME LAC.

**Validation:** JME LAC REN TRF KLT ANSM.

**Visualization:** LAC JME REN TRF.

**Writing – original draft:** JME LAC TRF REN.

**Writing – review & editing:** JME LAC ANSM KJA KLT.

## References

1. Rider LG, Nistala K. The juvenile idiopathic inflammatory myopathies: pathogenesis, clinical and auto-antibody phenotypes, and outcomes. *J Intern Med*. 2016; 280:24–38. doi: [10.1111/joim.12444](https://doi.org/10.1111/joim.12444) PMID: [27028907](https://pubmed.ncbi.nlm.nih.gov/27028907/)
2. Mendez EP, Lipton R, Ramsey-Goldman R, Roettcher P, Bowyer S, Dyer A, et al. US incidence of juvenile dermatomyositis, 1995–1998: results from the National Institute of Arthritis and Musculoskeletal and Skin Diseases Registry. *Arthritis Rheum*. 2003; 49:300–305. doi: [10.1002/art.11122](https://doi.org/10.1002/art.11122) PMID: [12794783](https://pubmed.ncbi.nlm.nih.gov/12794783/)
3. Bowyer SL, Blane CE, Sullivan DB, Cassidy JT. Childhood dermatomyositis: factors predicting functional outcome and development of dystrophic calcification. *J Pediatr*. 1983; 103:882–888. PMID: [6644423](https://pubmed.ncbi.nlm.nih.gov/6644423/)
4. Manlhiot C, Liang L, Tran D, Bitnun A, Tyrrell PN, Feldman BM. Assessment of an infectious disease history preceding juvenile dermatomyositis symptom onset. *Rheumatology* 2008; 47:526–529. doi: [10.1093/rheumatology/ken038](https://doi.org/10.1093/rheumatology/ken038) PMID: [18304939](https://pubmed.ncbi.nlm.nih.gov/18304939/)
5. Reed AM, Stirling JD. Association of the HLA-DQA1\*0501 allele in multiple racial groups with juvenile dermatomyositis. *Hum Immunol*. 1995; 44:131–135. PMID: [8666549](https://pubmed.ncbi.nlm.nih.gov/8666549/)
6. Pachman LM, Liotta-Davis MR, Hong DK, Kinsella TR, Mendez EP, Kinder JM, et al. TNFα-308A allele in juvenile dermatomyositis: Association with increased production of tumor necrosis factor α, disease duration, and pathologic calcifications. *Arthritis Rheum*. 2000; 43:2368–2377. doi: [10.1002/1529-0131\(200010\)43:10<2368::AID-ANR26>3.0.CO;2-8](https://doi.org/10.1002/1529-0131(200010)43:10<2368::AID-ANR26>3.0.CO;2-8) PMID: [11037898](https://pubmed.ncbi.nlm.nih.gov/11037898/)
7. Mamyrova G, O'Hanlon TP, Monroe JB, Carrick DM, Malley JD, Adams S, et al. Immunogenetic risk and protective factors for juvenile dermatomyositis in Caucasians. *Arthritis Rheum*. 2006; 54:3979–3987. doi: [10.1002/art.22216](https://doi.org/10.1002/art.22216) PMID: [17133612](https://pubmed.ncbi.nlm.nih.gov/17133612/)
8. Mamyrova G, O'Hanlon TP, Sillers L, Malley K, James-Newton L, Parks CGet al. Cytokine gene polymorphisms as risk and severity factors for juvenile dermatomyositis. *Arthritis Rheum*. 2008; 58:3941–3950. doi: [10.1002/art.24039](https://doi.org/10.1002/art.24039) PMID: [19035492](https://pubmed.ncbi.nlm.nih.gov/19035492/)
9. Miller FW, Cooper RG, Vencovsky J, Rider LG, Danko K, Wedderburn LR, et al. Genome-wide association study of dermatomyositis reveals genetic overlap with other autoimmune disorders. *Arthritis Rheum*. 2013; 65:3239–3247. doi: [10.1002/art.38137](https://doi.org/10.1002/art.38137) PMID: [23983088](https://pubmed.ncbi.nlm.nih.gov/23983088/)
10. Miller FW, Chen W, O'Hanlon TP, Cooper RG, Vencovsky J, Rider LG, et al. Genome-wide association study identifies HLA 8.1 ancestral haplotype alleles as major genetic risk factors for myositis phenotypes. *Genes Immun*. 2015; 16:470–480. doi: [10.1038/gene.2015.28](https://doi.org/10.1038/gene.2015.28) PMID: [26291516](https://pubmed.ncbi.nlm.nih.gov/26291516/)
11. Hargis AM, Haupt KH, Hegreberg GA, Prieur DJ, Moore MP. Familial canine dermatomyositis. Initial characterization of the cutaneous and muscular lesions. *Am J Pathol*. 1984; 116:234–244. PMID: [6465285](https://pubmed.ncbi.nlm.nih.gov/6465285/)
12. Hargis AM, Haupt KH, Prieur DJ, Moore MP. Animal model of human disease. Dermatomyositis. Familial canine dermatomyositis. *Am J Pathol*. 1985; 120:323–325. PMID: [4025513](https://pubmed.ncbi.nlm.nih.gov/4025513/)
13. Haupt KH, Prieur DJ, Hargis AM, Cowell RL, McDonald TL, Werner LL, et al. Familial canine dermatomyositis: clinicopathologic, immunologic, and serologic studies. *Am J Vet Res*. 1985; 46:1870–1875. PMID: [3901838](https://pubmed.ncbi.nlm.nih.gov/3901838/)
14. Haupt KH, Prieur DJ, Moore MP, Hargis AM, Hegreberg GA, Gavin PR, et al. Familial canine dermatomyositis: clinical, electrodiagnostic, and genetic studies. *Am J Vet Res*. 1985; 46:1861–1869. PMID: [4051292](https://pubmed.ncbi.nlm.nih.gov/4051292/)
15. Nagaraju K, Plotz PH. Animal models of myositis. *Rheum Dis Clin N Am*. 2002; 28:917–933.
16. Avci P, Sadasivam M, Gupta A, De Melo WC, Huang YY, Yin R, et al. Animal models of skin disease for drug discovery. *Expert Opin Drug Discov*. 2013; 8:331–355. doi: [10.1517/17460441.2013.761202](https://doi.org/10.1517/17460441.2013.761202) PMID: [23293893](https://pubmed.ncbi.nlm.nih.gov/23293893/)
17. Kunkle GA, Chrisman CL, Gross TL, Fadok V, Werner LL. Dermatomyositis in collie dogs. *Comp Contin Educ Pract Vet*. 1985; 7:185–192.
18. Hargis A, Mundell A. Familial canine dermatomyositis. *Compend Contin Educ Vet*. 1992; 4:855–864.
19. Scott DW, Miller WH, Griffin CE. Congenital and hereditary defects. In: Muller and Kirk's Small Animal Dermatology, 6th ed. Philadelphia: Saunders, 2000 pp. 940–946.

20. Wahl JM, Clark LA, Skalli O, Ambrus A, Rees CA, Mansell JL, et al. Analysis of gene transcript profiling and immunobiology in Shetland sheepdogs with dermatomyositis. *Vet Dermatol*. 2008; 19:52–58. doi: [10.1111/j.1365-3164.2008.00655.x](https://doi.org/10.1111/j.1365-3164.2008.00655.x) PMID: [18336421](https://pubmed.ncbi.nlm.nih.gov/18336421/)
21. Hargis AM, Prieur DJ, Haupt KH, McDonald TL, Moore MP. Prospective study of familial canine dermatomyositis: correlation of the severity of dermatomyositis and circulating immune complex levels. *Am J Pathol*. 1986; 123:465–479. PMID: [3717300](https://pubmed.ncbi.nlm.nih.gov/3717300/)
22. Rider LG, Wu L, Mamyrova G, Targoff IN, Miller FW. Environmental factors preceding illness onset differ in phenotypes of the juvenile idiopathic inflammatory myopathies. *Rheumatology* 2010; 49:2381–2390. doi: [10.1093/rheumatology/keq277](https://doi.org/10.1093/rheumatology/keq277) PMID: [20802007](https://pubmed.ncbi.nlm.nih.gov/20802007/)
23. Dourmishev LA, Dourmishev AL. *Dermatomyositis: Advances in Recognition, Understanding and Management*. New York: Springer; 2009.
24. vonHoldt BM, Pollinger JP, Lohmueller KE, Han E, Parker HG, Quignon P, et al. Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. *Nature* 2010; 464:898–902. doi: [10.1038/nature08837](https://doi.org/10.1038/nature08837) PMID: [20237475](https://pubmed.ncbi.nlm.nih.gov/20237475/)
25. Wilbe M, Jokinen P, Truvé K, Seppala EH, Karlsson EK, Biagi T et al. Genome-wide association mapping identifies multiple loci for a canine SLE-related disease complex. *Nat Genet*. 2010; 42:250–254. doi: [10.1038/ng.525](https://doi.org/10.1038/ng.525) PMID: [20101241](https://pubmed.ncbi.nlm.nih.gov/20101241/)
26. Karlsson EK, Sigurdsson S, Ivansson E, Thomas R, Elvers I, Wright J et al. Genome-wide analyses implicate 33 loci in heritable dog osteosarcoma, including regulatory variants near CDKN2A/B. *Genome Biol*. 2013; 14:R132. doi: [10.1186/gb-2013-14-12-r132](https://doi.org/10.1186/gb-2013-14-12-r132) PMID: [24330828](https://pubmed.ncbi.nlm.nih.gov/24330828/)
27. Karyadi DM, Karlins E, Decker B, vonHoldt BM, Carpintero-Ramirez G, Parker HG et al. A copy number variant at the KITLG locus likely confers risk for canine squamous cell carcinoma of the digit. *PLoS Genet*. 2013; 9:e1003409. doi: [10.1371/journal.pgen.1003409](https://doi.org/10.1371/journal.pgen.1003409) PMID: [23555311](https://pubmed.ncbi.nlm.nih.gov/23555311/)
28. Tang R, Noh HJ, Wang D, Sigurdsson S, Swofford R, Perloski M, et al. Candidate genes and functional noncoding variants identified in a canine model of obsessive-compulsive disorder. *Genome Biol*. 2014; 15:R25. doi: [10.1186/gb-2014-15-3-r25](https://doi.org/10.1186/gb-2014-15-3-r25) PMID: [24995881](https://pubmed.ncbi.nlm.nih.gov/24995881/)
29. Ivansson EL, Megquier K, Kozyrev SV, Murén E, Körberg IB, Swofford R, et al. Variants within the SP110 nuclear body protein modify risk of canine degenerative myelopathy. *Proc Natl Acad Sci*. 2016; 113:E3091–100. doi: [10.1073/pnas.1600084113](https://doi.org/10.1073/pnas.1600084113) PMID: [27185954](https://pubmed.ncbi.nlm.nih.gov/27185954/)
30. Melin M, Rivera P, Arendt M, Elvers I, Murén E, Gustafson U, et al. Genome-wide analysis identifies germ-line risk factors associated with canine mammary tumours. *PLoS Genet*. 2016; 12:e1006029. doi: [10.1371/journal.pgen.1006029](https://doi.org/10.1371/journal.pgen.1006029) PMID: [27158822](https://pubmed.ncbi.nlm.nih.gov/27158822/)
31. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc*. 2009; 4:1073–1081. doi: [10.1038/nprot.2009.86](https://doi.org/10.1038/nprot.2009.86) PMID: [19561590](https://pubmed.ncbi.nlm.nih.gov/19561590/)
32. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nat Methods* 2010; 7:248–249. doi: [10.1038/nmeth0410-248](https://doi.org/10.1038/nmeth0410-248) PMID: [20354512](https://pubmed.ncbi.nlm.nih.gov/20354512/)
33. Tang H, Thomas PD. PANTHER-PSEP: predicting disease-causing genetic variants using position-specific evolutionary preservation. *Bioinformatics*. 2016; 32:2230–2232. doi: [10.1093/bioinformatics/btw222](https://doi.org/10.1093/bioinformatics/btw222) PMID: [27193693](https://pubmed.ncbi.nlm.nih.gov/27193693/)
34. Tucker M, Valencia-Sanchez MA, Staples RR, Chen J, Denis CL, Parker R. The transcription factor associated Ccr4 and Caf1 proteins are components of the major cytoplasmic mRNA deadenylase in *Saccharomyces cerevisiae*. *Cell* 2001; 104:377–386. PMID: [11239395](https://pubmed.ncbi.nlm.nih.gov/11239395/)
35. Wolf J, Passmore LA. mRNA deadenylation by Pan2–Pan3. *Biochem Soc Trans*. 2014; 42:184–187. doi: [10.1042/BST20130211](https://doi.org/10.1042/BST20130211) PMID: [24450649](https://pubmed.ncbi.nlm.nih.gov/24450649/)
36. Weill L, Belloc E, Bava FA, Méndez R. Translational control by changes in poly(A) tail length: recycling mRNAs. *Nat Struct Mol Biol*. 2012; 19:577–585. doi: [10.1038/nsmb.2311](https://doi.org/10.1038/nsmb.2311) PMID: [22664985](https://pubmed.ncbi.nlm.nih.gov/22664985/)
37. Bett JS, Ibrahim AF, Garg AK, Kelly V, Pedrioli P, Rocha S, et al. The P-body component USP52/PAN2 is a novel regulator of HIF1A mRNA stability. *Biochem J*. 2013; 451:185–194. doi: [10.1042/BJ20130026](https://doi.org/10.1042/BJ20130026) PMID: [23398456](https://pubmed.ncbi.nlm.nih.gov/23398456/)
38. Palazon A, Goldrath AW, Nizet V, Johnson RS. HIF transcription factors, inflammation, and immunity. *Immunity* 2014; 41:518–528. doi: [10.1016/j.immuni.2014.09.008](https://doi.org/10.1016/j.immuni.2014.09.008) PMID: [25367569](https://pubmed.ncbi.nlm.nih.gov/25367569/)
39. Bakheet T, Frevel M, Williams BR, Greer W, Khabar KS. ARED: human AU-rich element-containing mRNA database reveals an unexpectedly diverse functional repertoire of encoded proteins. *Nucleic Acids Res*. 2001; 29:246–254. PMID: [11125104](https://pubmed.ncbi.nlm.nih.gov/11125104/)
40. Preuße C, Allenbach Y, Hoffmann O, Goebel HH, Pehl D, Radke J, et al. Differential roles of hypoxia and innate immunity in juvenile and adult dermatomyositis. *Acta Neuropathol Commun*. 2016; 4:45. doi: [10.1186/s40478-016-0308-5](https://doi.org/10.1186/s40478-016-0308-5) PMID: [27121733](https://pubmed.ncbi.nlm.nih.gov/27121733/)



41. Bilgic H, Ytterberg SR, Amin S, McNallan KT, Wilson JC, Koeuth T, et al. Interleukin-6 and type I interferon-regulated genes and chemokines mark disease activity in dermatomyositis. *Arthritis Rheum*. 2009; 60:3436–3446. doi: [10.1002/art.24936](https://doi.org/10.1002/art.24936) PMID: [19877033](https://pubmed.ncbi.nlm.nih.gov/19877033/)
42. Li J, Ji C, Yang Q, Chen J, Gu S, Ying K, et al. Cloning and characterization of a novel human TGF-beta activated kinase-like gene. *Biochem Genet*. 2004; 42:129–137. PMID: [15168726](https://pubmed.ncbi.nlm.nih.gov/15168726/)
43. Kilpinen S, Ojala K, Kallioniemi O. Analysis of kinase gene expression patterns across 5681 human tissue samples reveals functional genomic taxonomy of the kinome. *PLoS One* 2010; 5:e15068. doi: [10.1371/journal.pone.0015068](https://doi.org/10.1371/journal.pone.0015068) PMID: [21151926](https://pubmed.ncbi.nlm.nih.gov/21151926/)
44. Gupta S, Stamatoyannopoulos JA, Bailey TL, Noble WS. Quantifying similarity between motifs. *Genome Biol*. 2007; 8:R24. doi: [10.1186/gb-2007-8-2-r24](https://doi.org/10.1186/gb-2007-8-2-r24) PMID: [17324271](https://pubmed.ncbi.nlm.nih.gov/17324271/)
45. Woolf E, Xiao C, Fainaru O, Lotem J, Rosen D, Negreanu V, et al. Runx3 and Runx1 are required for CD8 T cell development during thymopoiesis. *Proc Natl Acad Sci USA*. 2003; 100:7731–7736. doi: [10.1073/pnas.1232420100](https://doi.org/10.1073/pnas.1232420100) PMID: [12796513](https://pubmed.ncbi.nlm.nih.gov/12796513/)
46. Djuretic IM1, Cruz-Guilloty F, Rao A. Regulation of gene expression in peripheral T cells by Runx transcription factors. *Adv Immunol*. 2009; 104:1–23. doi: [10.1016/S0065-2776\(08\)04001-7](https://doi.org/10.1016/S0065-2776(08)04001-7) PMID: [20457114](https://pubmed.ncbi.nlm.nih.gov/20457114/)
47. Fainaru O, Woolf E, Lotem J, Yarmus M, Brenner O, Goldenberg D, et al. Runx3 regulates mouse TGF-beta-mediated dendritic cell function and its absence results in airway inflammation. *EMBO J*. 2004; 23:969–979. doi: [10.1038/sj.emboj.7600085](https://doi.org/10.1038/sj.emboj.7600085) PMID: [14765120](https://pubmed.ncbi.nlm.nih.gov/14765120/)
48. Brenner O, Levanon D, Negreanu V, Golubkov O, Fainaru O, Woolf E, et al. Loss of Runx3 function in leukocytes is associated with spontaneously developed colitis and gastric mucosal hyperplasia. *Proc Natl Acad Sci USA*. 2004; 101:16016–16021. doi: [10.1073/pnas.0407180101](https://doi.org/10.1073/pnas.0407180101) PMID: [15514019](https://pubmed.ncbi.nlm.nih.gov/15514019/)
49. Guo C, Yao F, Wu K, Yang L, Zhang X, Ding J. Chromatin immunoprecipitation and association study revealed a possible role of Runt-related transcription factor 3 in the ulcerative colitis of Chinese population. *Clin Immunol*. 2010; 135:483–489. doi: [10.1016/j.clim.2010.01.004](https://doi.org/10.1016/j.clim.2010.01.004) PMID: [20392673](https://pubmed.ncbi.nlm.nih.gov/20392673/)
50. Evans DM, Spencer CC, Pointon JJ, Su Z, Harvey D, Kochan G, et al. Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility. *Nat Genet*. 2011; 43:761–767. doi: [10.1038/ng.873](https://doi.org/10.1038/ng.873) PMID: [21743469](https://pubmed.ncbi.nlm.nih.gov/21743469/)
51. Helms C, Cao L, Krueger JG, Wijsman EM, Chamian F, Gordon D, et al. A putative RUNX1 binding site variant between SLC9A3R1 and NAT9 is associated with susceptibility to psoriasis. *Nat Genet*. 2003; 35:349–356. doi: [10.1038/ng1268](https://doi.org/10.1038/ng1268) PMID: [14608357](https://pubmed.ncbi.nlm.nih.gov/14608357/)
52. Prokunina L, Castillejo-López C, Oberg F, Gunnarsson I, Berg L, Magnusson V, et al. A regulatory polymorphism in PDCD1 is associated with susceptibility to systemic lupus erythematosus in humans. *Nat Genet*. 2002; 32:666–669. doi: [10.1038/ng1020](https://doi.org/10.1038/ng1020) PMID: [12402038](https://pubmed.ncbi.nlm.nih.gov/12402038/)
53. Hallgrímsdóttir IB, Yuster DS. A complete classification of epistatic two-locus models. *BMC Genet*. 2008; 9:17. doi: [10.1186/1471-2156-9-17](https://doi.org/10.1186/1471-2156-9-17) PMID: [18284682](https://pubmed.ncbi.nlm.nih.gov/18284682/)
54. Mackay TR. Epistasis and quantitative traits: using model organisms to study gene-gene interactions. *Nat Rev Genet*. 2014; 15:22–33. doi: [10.1038/nrg3627](https://doi.org/10.1038/nrg3627) PMID: [24296533](https://pubmed.ncbi.nlm.nih.gov/24296533/)
55. Parker WM, Foster RA. Cutaneous vasculitis in five Jack Russell Terriers. *Vet Dermatol*. 1996; 7:109–115.
56. Ihrke PJ. Ischemic skin disease in the dog. 31st World Small Animal Veterinary Congress (2006).
57. Kennedy LJ, Barnes A, Short A, Brown JJ, Lester S, Seddon J, et al. Canine DLA diversity: 1. New alleles and haplotypes. *Tissue Antigens* 2007; 69:272–288. doi: [10.1111/j.1399-0039.2006.00779.x](https://doi.org/10.1111/j.1399-0039.2006.00779.x) PMID: [17445218](https://pubmed.ncbi.nlm.nih.gov/17445218/)
58. Clark LA, Wahl JM, Rees CA, Murphy KE. Retrotransposon insertion in SILV is responsible for merle patterning of the domestic dog. *Proc Natl Acad Sci USA* 2006; 103:1376–1381. doi: [10.1073/pnas.0506940103](https://doi.org/10.1073/pnas.0506940103) PMID: [16407134](https://pubmed.ncbi.nlm.nih.gov/16407134/)
59. Evans JM, Cox ML, Huska J, Li F, Gaitero L, Guo LT, et al. Exome sequencing reveals a nebulin non-sense mutation in a dog model of nemaline myopathy. *Mamm Genome* 2016; 27:495–502. doi: [10.1007/s00335-016-9644-9](https://doi.org/10.1007/s00335-016-9644-9) PMID: [27215641](https://pubmed.ncbi.nlm.nih.gov/27215641/)
60. Kennedy LJ, Davison LJ, Barnes A, Short AD, Fretwell N, Jones CA, et al. Identification of susceptibility and protective major histocompatibility complex haplotypes in canine diabetes mellitus. *Tissue Antigens* 2006; 68:467–476. doi: [10.1111/j.1399-0039.2006.00716.x](https://doi.org/10.1111/j.1399-0039.2006.00716.x) PMID: [17176436](https://pubmed.ncbi.nlm.nih.gov/17176436/)
61. Robinson J, Halliwell JA, McWilliam H, Lopez R, Marsh SGE. IPD—the immuno polymorphism database. *Nucleic Acids Res*. 2013; 41:1234–1240.
62. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP et al. LocusZoom: Regional visualization of genome-wide association scan results. *Bioinformatics* 2010; 26:2336–7. doi: [10.1093/bioinformatics/btq419](https://doi.org/10.1093/bioinformatics/btq419) PMID: [20634204](https://pubmed.ncbi.nlm.nih.gov/20634204/)



63. Langmead B, Salzberg S. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012; 9:357–359. doi: [10.1038/nmeth.1923](https://doi.org/10.1038/nmeth.1923) PMID: [22388286](https://pubmed.ncbi.nlm.nih.gov/22388286/)
64. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence alignment/map (SAM) format and SAMtools. *Bioinformatics* 2009; 25:2078–2079. doi: [10.1093/bioinformatics/btp352](https://doi.org/10.1093/bioinformatics/btp352) PMID: [19505943](https://pubmed.ncbi.nlm.nih.gov/19505943/)
65. Li H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 2011; 27:2987–2993. doi: [10.1093/bioinformatics/btr509](https://doi.org/10.1093/bioinformatics/btr509) PMID: [21903627](https://pubmed.ncbi.nlm.nih.gov/21903627/)
66. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative Genomics Viewer. *Nat Biotechnol*. 2011; 29:24–26. doi: [10.1038/nbt.1754](https://doi.org/10.1038/nbt.1754) PMID: [21221095](https://pubmed.ncbi.nlm.nih.gov/21221095/)
67. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014; 30:2114–2120. doi: [10.1093/bioinformatics/btu170](https://doi.org/10.1093/bioinformatics/btu170) PMID: [24695404](https://pubmed.ncbi.nlm.nih.gov/24695404/)
68. Wu TD, Nacu S. Fast and SNP-tolerant detection of complex variants and splicing in short reads. *Bioinformatics* 2010; 26:873–881. doi: [10.1093/bioinformatics/btq057](https://doi.org/10.1093/bioinformatics/btq057) PMID: [20147302](https://pubmed.ncbi.nlm.nih.gov/20147302/)
69. Klein JP, Moeschberger ML. *Survival analysis: techniques for censored and truncated data*. New York: Springer; 2005.
70. Gelman A. Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Anal*. 2006; 1:515–534.
71. Carpenter B, Gelman A, Hoffman M, Lee D, Goodrich B, Betancourt M, et al. Stan: A probabilistic programming language. *J Stat Softw*. 2016;Forthcoming
72. Gelman A, Rubin D. Inference from iterative simulation using multiple sequences. *Statist Sci*. 1992; 7:457–472.
73. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. *Bayesian Data Analysis*, 3rd ed. Broken Sound Parkway: CRC Press Taylor & Francis Group; 2014.
74. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria; 2015.
75. Qanbari S, Strom TM, Haberer G, Weigend S, Gheyas AA, Turner F, et al. A high resolution genome-wide scan for significant selective sweeps: an application to pooled sequence data in laying chickens. *PLoS One* 2012; 7:e49525. doi: [10.1371/journal.pone.0049525](https://doi.org/10.1371/journal.pone.0049525) PMID: [23209582](https://pubmed.ncbi.nlm.nih.gov/23209582/)
76. Warnes G, Gorjanc G, Leisch F, Man M. *Genetics: Population Genetics*. R package version 1.3.8.1 2013.